APPLICATION OF STATISTICAL QUALITY CONTROL PROCEDURES TO PRODUCTION OF HIGHWAY PAVEMENT CONCRETE

JANUARY 1966
NO. 3

by

S. J. HANNA
J. F. McLAUGHLIN
A. P. LOTT

Joint Highway Research Project

PURDUE UNIVERSITY
LAFAYETTE INDIANA

Technical Paper

APPLICATION OF STATISTICAL QUALITY CONTROL PROCEDURES TO PRODUCTION OF HIGHWAY PAVEMENT CONCRETE

To: G. A. Leonards, Director
Joint Highway Research Project

February 11, 1966

From: H. L. Michael, Associate Director
Joint Highway Research Project

File: 9-11-2
Project: C-36-67B

Attached is a paper entitled "Application of Statistical Quality Control Procedures to Production of Highway Pavement Concrete" co-authored by S. J. Hanna, J. F. McLaughlin and A. P. Lott.

This work was presented orally at the last meeting of the Highway Research Board. It is a summary of the first phase of the quality control project previously reported to the Board. Permission to publish this report in the HRB Research Record is requested.

Respectfully submitted,

H. L. Michael /JFM

H. L. Michael, Secretary

HLM:kr

Attachment

cc: F. L. Ashbaucher
J. R. Cooper
J. W. Delleur
W. L. Dolch
W. H. Goetz
W. L. Grecco
G. K. Hallock
F. S. Hill
J. F. McLaughlin
F. B. Mendenhall
R. D. Miles
J. C. Oppenlander
W. P. Privette
M. B. Scott
F. W. Stubbs
K. B. Woods
E. J. Yoder

Technical Paper

APPLICATION OF STATISTICAL QUALITY CONTROL PROCEDURES TO PRODUCTION OF HIGHWAY PAVEMENT CONCRETE

by

S. J. Hanna, Research Assistant
J. F. McLaughlin, Research Engineer
A. P. Lott, Graduate Assistant

Joint Highway Research Project

Project: C-36-67B

File: 9-11-2

Prepared as Part of an Investigation

Conducted by

Joint Highway Research Project
Engineering Experiment Station
Purdue University

in cooperation with

Indiana State Highway Commission

and the

Bureau of Public Roads
U.S. Department of Commerce

Not Released for Publication

Subject to Change

Not Reviewed By
Indiana State Highway Commission
or the
Bureau of Public Roads

Purdue University
Lafayette, Indiana
January 1966

INTRODUCTION

This investigation was concerned with the collection of data by a systematic procedure for the purpose of evaluating the variability present in the manufacture of portland cement concrete for highway pavements. The data were analyzed to provide information concerning the magnitude of the variance components for the Bureau of Public Roads' data system and to provide information and illustrate procedures for the establishment of a quality control program that could be used by the Indiana State Highway Commission.

Over the years many specifications have been evolved through trial and error without reference to the actual variability of the product or process. In theory it is possible to improve the product by narrowing the specification limits, but if the process itself is incapable of operating within those limits then they are of little use. It is, as has been stated, one of the aims of this investigation to obtain estimates of the variability associated with the manufacturing of fresh portland cement concrete for highway pavement.

Specification requirements are of little use unless some means of testing and control are exerted. With estimates of the variability at hand, it is possible to develop a quality control program based on a thorough understanding of the capabilities of the process. Also, it is possible to establish a realistic system and schedule of acceptance tests, number of samples, etc.

The construction of a highway may be likened to an industrial manufacturing process. There is a manufactured product, the highway, and like industrial production there is a need to control the quality of the product. This need arises from the desire of the manufacturer, the contractor, to produce a product for the purchaser, the State, in the most economical manner possible while meeting the specifications for the product. The purchaser in turn is interested in seeing that he obtains a quality product.

Statistical quality control provides a means whereby a manufacturer can derive maximum benefit from control testing of the manufactured product. The basic concepts are applicable whether the product be piston rings or highway pavements. Inherent in statistical analyses is the ability to make estimates of population parameters from sample statistics and to associate with these estimates of the probability of being in error. Using statistical quality control procedures, a manufacturing process can be investigated to determine the range in values that one can expect under existing conditions. This information is valuable to the producer and to the purchaser. It can be used not only in determining compliance with specifications but also to judge whether the construction or manufacturing process is capable of producing the product within them. If existing specifications are unrealistic with respect to an end result or are economically unattainable, quality control data can provide a basis for the development of revised standards.

OUTLINE OF WORK

Plastic portland cement concrete was chosen as the area of investigation. The specific area was limited to concrete paving projects under contract in Indiana, and tests for air content, slump and unit weight were made on the concrete. Air content was determined using both the pressure type air meter and the Chace air meter. These tests were conducted by a research team from Purdue University and all tests were made independent of Indiana State Highway Commission control tests.

Three paving projects were selected in cooperation with the Indiana State Highway Commission, with each project performed by a different contractor. The projects were chosen on the basis of their geographic location in the state and the paving schedules of the contractors.

Three replicate determinations of each attribute (slump, air content and unit weight) were made on fifty samples obtained on each project. Hence for this investigation 150 individual tests were performed for each test method on all projects for a total of 450 observations over the three projects. The replicate determinations were selected rather than two samples tested twice from each location because of the time involved in making a test and the number of different tests being performed.

On each paving project sampling began at the start of paving operations for any one day by the random selection of a batch and then continued throughout the day at time intervals dictated by the time required for each set-up. It is considered that this provided a random procedure that eliminated bias in the sampling procedure. The time for each set-up varied considerably because of variations in the distance from sampling point, and ease of movement of equipment. A typical set-up from start to finish required approximately one hour.

The data were collected during the summer construction season of 1964. The raw data were placed on IBM punch cards with appropriate coding to indicate job number, sample number, replicate number, time of test and date test was made. The data were analyzed using standard statistical techniques and procedures. The IBM 7094 computer was utilized in the data analysis.

FIELD PROCEDURES AND TESTING

After the four tests had been selected (air content by both pressure meter and Chace meters, slump and unit weight) equipment and personnel were organized. It was quite obvious at the outset that the whole operation had to be a highly mobile one. The equipment had to be transported to each of the three projects and then moved along the paving operation from test point to test point. It was felt the best way to handle the problem was through the use of a pick-up truck. The truck was outfitted with a few attachments to facilitate the testing program. A plywood box was bolted to the body of the truck and used for storage of various small items such as tamping rods, trowels, etc. It was also noted that the testing program would be such that it would be necessary to have a supply of water on hand at all times. A 55 gallon drum with a hose and spiget attached was strapped to the left side of the truck body. This drum proved to be quite handy and made the operation extremely self sufficient.

With the equipment and vehicle in order, job sites were selected. As mentioned previously, each site selected was selected on the basis of geographic location in the state and on the basis of their paving schedules. (Since the testing program was limited to the summer months of 1964 only sites with paving in progress were considered). As soon as a site was selected, a team of operators went to the site to begin the testing program. The teams consisted of two men for the first site and a part of the second but was expanded to three men for the remainder of the second site and all of the third. The two persons doing the actual testing were never changed, and they performed the same tests throughout the whole research project. Operator A performed the slump and unit weight tests while Operator B performed both types of air content tests.

The site was surveyed to determine where and how to begin the testing program. Also, pertinent information was obtained concerning the mix design, sources and types of materials, any correction factors and other data needed for the testing.

The testing of a single sample of concrete required anywhere from 30 minutes to an hour and fifteen minutes from start of sampling to final cleanup. Four different tests were performed in triplicate on each sample so there was little time to waste before the concrete would begin to stiffen. After some experience, this procedure became a highly efficient operation.

All the testing was performed on the right side of the forms in the direction of pouring. The dual-drum pavers and auxiliary equipment were located on the median side and a set-up there would mean disturbing the concreting operations. The one guiding principal was to stay completely out of the way of the paving operations. Working on the right shoulder created one problem in that this was where the contractor normally laid out his steel. In some cases this meant a longer distance from sampling point to where the equipment was set or, where the subbase was especially wide, working to the right of the steel.

The set-up for the testing was placed as close to the forms as was possible without interference. The set-up took about 5 minutes and required placing three square pieces of plywood and positioning the testing equipment. The plywood served as working platforms for the scale, slump tests and air tests.

Concrete was sampled from the batch which had been deposited on the grade. The sample of fresh concrete was placed in a wheelbarrow and a large pan. Approximately three cubic feet of concrete were required for each sample. The sample was obtained before the batch was spread by the first spreader in the case of an operation using twin-barrel mixes and after the initial spread in the case of a central mix operation. The distance between samples was quite arbitrary and depended upon how far the paving train progressed between set-ups and how long it took the team to perform the tests. The sampling operation required a maximum of 5 minutes.

With the concrete sample having been obtained, the tests themselves were performed. Both Operators A and B started simultaneously performing their respective tests. The equipment was positioned so the testing could begin immediately to provide the maximum amount of time before the concrete began to stiffen. Operator B immediately started performing the air content test by the pressure method while Operator A started on the slump tests. These tests were performed in accordance with ASTM standards.

ANALYSIS OF DATA

At the completion of the testing program all data were tabulated and recorded on IBM punch cards. Information regarding job number, sample number, replicate number, time of test and date was coded and placed on the punch cards along with the appropriate data for ease of identification. The statistical analysis of the data was accomplished using standard computer programs for analysis of variance, correlation and distribution. In addition, standard statistical techniques and procedures were utilized to determine confidence limits, control limits and in significance testing. A majority of the analyses and plotting of data was accomplished using the IBM 7094-1410 computer system.

The data collected from each of the four tests (air content by pressure meter, air content by Chace meter, slump and unit weight) were analyzed separately and the sum of squares, mean squares and standard deviations computed for each test method. The first analysis was based upon a 2-factor factorial design model with three replicate observations for one factor (samples). In addition, correlation coefficients were determined for all combinations of the above mentioned tests. Sample means were used in the correlations and data plotting.

In the development of a quality control program it is necessary to obtain data from a process which is "in control," that is, from a process in which the variability is due to chance causes alone and not to assignable causes. From observations in the field, such as noting obvious errors in air-entraining agent content, water content, etc. it can be said that at certain times a portion of the variability noted in the present investigation was due to assignable errors. For this reason a one-way analysis of variance was conducted for each site separately in addition to the factorial analysis.

In certain of the analyses it was noted that the magnitude of the variance components differed from site to site. Analyzing the data for each site separately allows the computation of these variance components and makes it possible to compare the magnitude of the components from site to site. A factorial analysis averages the variances from the three sites and hence if at one or two sites the process is out of control, there is no estimate available for the variance of an in control process. In fact the factorial analysis is invalid if the variances are not homogeneous (i.e., variances are not statistically equal).

The factorial analyses have been included in this report for the purpose of illustrating this type of statistical procedure. If other variables such as operator or equipment were included in an investigation the factorial design model could be used in the analysis of the data.

It should be noted that operators and testing equipment were not considered as variables in this investigation. Only one operator and one piece of testing equipment was used throughout the investigation for each test method. This necessarily limits the interpretation of the data. The values of standard deviations and confidence limits cannot be applied directly to a project on which several operators and several pieces of testing equipment are used.

As a sample was tested in the field for air content by the pressure meter, a time dependency was observed. This led to testing the differences between replicates and calculation of the correlation coefficient associated with the third pressure replicate versus the sample mean of the Chace tests. Results of this phase of the investigation will be discussed in a later section.

The test results were also used to illustrate techniques and procedures that may be employed in a quality control program. Control limits are illustrated in the section on Quality Control.

For simplicity and ease in handling the large amount of data, a discussion of each test method will be presented separately. Sections concerning correlations and quality control applications follow. A summary of a portion of the basic statistical results is presented in the Appendix.

Field Observations

Dual-drum pavers were used on Sites 1 and 3 while a central mix plant was in operation on Site 2. These were quite different sets of conditions depending on the type of paving operation being employed. The basic difference between the sites was the method of mixing with all other operations being essentially the same.

Each method of paving had its own characteristics of control with respect to frequency of adjustment. Quite often with the dual-drum pavers the water valve was adjusted and readjusted to allow more or less water into each batch. This yielded many batches that were alternately wet or dry. This variability in water content per batch was due also to the use of dry and wet batches of aggregate.

In the central mix project there were fewer adjustments. The plant was started up and checked at the start of the project but then almost complete reliance was placed on the automatic features of the plant. Thus, there was less checking and less control of the concrete. The major problem was control of air content. By the time a low air content was noticed and a message relayed to the plant to make the necessary changes, many concrete trucks were either dumping or already on their way to the grade with their 8 cubic yards of concrete. There was a large lag-time between catching a low air reading and effecting a correction. This was an unfortunate characteristic of the operation.

It was noticed that the less the paving operation is changed, the more constant the concrete product. This was quite evident at Site 3 where very few adjustments were made in the way of water content, air entraining agent or batch changes. This fact is substantiated by the statistical analysis. Site 3 has the best grouping of data and distribution of results.

Air Content by Pressure Meter

The analysis of variance, hereafter referred to as the ANOV, for the air content measured by pressure meter is presented in Table 1. The sources of variation as determined by the factorial model are: site-to-site variation, sample-within-site variation and the error term. Table 2 presents a summary of the statistical analysis results based upon a factorial design model.

A standard test for significance, the F-test, indicates that at the 0.05 α-level (probability of rejecting the hypothesis when it is true) the site-to-site variation is not significant but the sample-within-site variation is significant. The concrete is manufactured in batches and a sample comes from a single batch, hence the sample-within-site variation is a measure of the batch-to-batch variation. Therefore, at an α-level of 0.05 the batch means are different.

When first viewed, these results may appear to be reversed from what one would expect. However, consider the manufacturing process. The sample-within-site variation is the batch-to-batch variation for a particular site. Changes in moisture content of the aggregate, adjustments in the amount of water per batch and adjustments in the amount of air-entraining agent can occur from batch-to-batch and one would expect the air content to change. Site-to-site variation would also be expected to be significant (different).

Table 1

Analysis of Variance - Factorial Model

Air Content by Pressure Test

Source of Variation	df	SS	MS	EMS	F
Site	2	8.19000	4.09500	$\sigma_\epsilon^2 + 3\sigma_s^2 + 150\ \sigma_{site}^2$	2.72*
Sample within Site	147	221.21985	1.50490	$\sigma_\epsilon^2 + 3\sigma_s^2$	19.09
Error	300	23.64664	0.07882	σ_ϵ^2	

$\alpha = 0.05$

$F_{2,147} = 3.07^* > 2.72$ $\therefore$ site not significant

$F_{147,300} = 1.30 < 19.09$ $\therefore$ sample within site significant

$$\sigma_s^2 = \frac{1.50490 - 0.07882}{3} = \frac{1.42608}{3} = 0.47536$$

TABLE 2

SUMMARY OF STATISTICAL ANALYSES

(FACTORIAL MODEL)

Standard Deviations (all sites)	Air Content Pressure	Air Content Chace	Slump	Unit Weight
Site Std. Dev.	0.10	0.11	0.14	0.15
Sample Mean Std. Dev.*	0.16	0.32	0.21	0.62
Sample-Within-Site Std. Dev.	0.69	0.69	0.98	0.89
Error Term (all sites)	0.079	0.30	0.14	1.15

* Consists of variation due to variance among determinations but not among samples.

It is necessary to understand the composition of the site-to-site variance, or in statistical terms, the expected mean square (EMS) components of variance. The EMS from Table 1 is $(\sigma_\epsilon^2 + 3\sigma^2_{sample} + 150\ \sigma^2_{site})$ for the site to site component; $(\sigma_\epsilon^2 + 3\sigma^2_{sample})$ for the sample-within-site component and σ_ϵ^2 for the error term. The error term (σ_ϵ^2) is observed to be small in comparison to the sample-within-site term $(\sigma_\epsilon^2 + 3\sigma_s^2)$ leading to the conclusion that sample-within-site variation is significant or that sample means are different. The σ_s^2 term is large compared to the σ^2_{site} term and when a significance test is performed: $(\sigma_\epsilon^2 + 3\sigma_s^2 + 150\ \sigma^2_{site})/\ (\sigma_\epsilon^2 + 3\sigma_s^2)$ the site-to-site component is determined not significant. If the distribution of sample means was smaller (i.e. σ_s^2 smaller) and σ_{site}^2 remained the same, a significance test might indicate the site-to-site component significant. In other words distribution of sample means is so large that it overshadows the spread among site means.

The distribution of air content for all sites measured by the pressure meter is shown in Figure 1. Values tabulated are sample means. The overall mean air content is 4.40 percent. The distribution over all sites approximates a normal distribution. The air content determinations for Sites 1 and 3 show some tendency towards normality but for Site 2 the distribution was definitely not normal. This may be accounted for by the fact that a number of difficulties arose with the plant operation on Site 2. The aggregate varied considerably in its moisture content and a number of failures occurred in the air entraining agent dispensing equipment. These factors combined to produce a large range in air contents and a non-normal distribution.

The observed error term from Table 2 is 0.079, or from a practical viewpoint 0.1%, indicating that an error of 0.1% can occur in the air

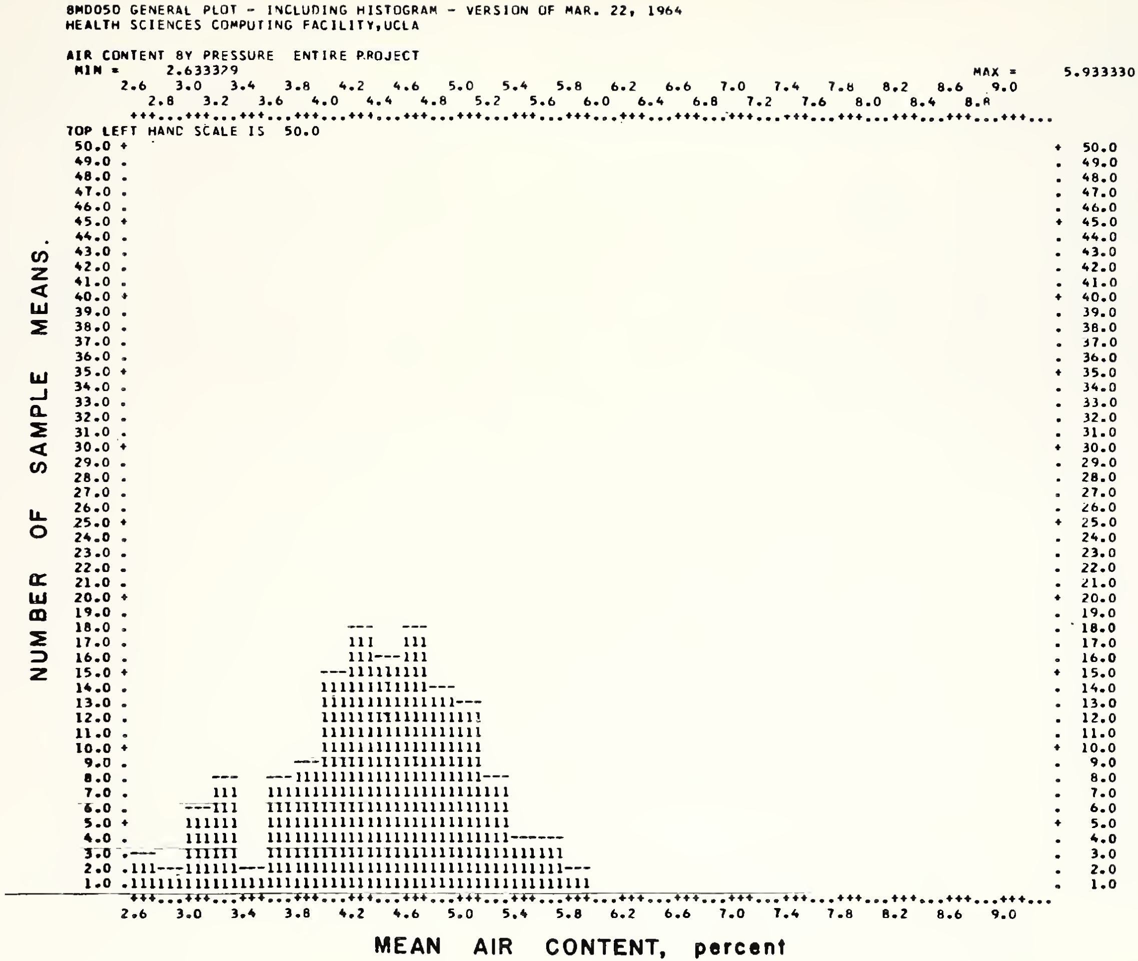

FIG. I HISTOGRAM FOR THE DISTRIBUTION OF AIR CONTENT MEASURED BY PRESSURE METER, ALL SITES.

content determination due to chance alone. Placing 95% confidence limits on the site mean gives a range within which we are 95% confident the true site mean lies. For example, the mean for Site 1 is 4.48%, therefore, we are 95% confident that the true site mean lies between 4.28% and 4.68%.

As mentioned previously, it was observed that assignable causes in several instances added to the measured variation and hence a one-way ANOV was performed on each test method for each site separately. A summary of the results are presented in Table 3.

If the mean square terms (MS) for the three sites as analyzed separately are averaged, the resulting average is equal to the corresponding mean square as determined by the ANOV of the factorial model. This provides a check as to the accuracy of the computation and illustrates how the mean square terms are related.

Note the differences in the Mean Square terms (MS) and the standard deviations from site to site.

Air Content by Chace Meter

The ANOV for air content by Chace meter is similar to that for air content by pressure meter (see Table 1). The statistical sources of variation are the same as those associated with the pressure meter. A summary of results from the statistical analyses is presented in Table 2. It should be noted that air contents by the Chace meter were determined in the field to the nearest one-half percent. Corrections for mortar content of the mix were computed and the appropriate adjustment made in the air content. The calculations in the statistical analysis portion of the investigation were carried to hundredths of percent for purposes of handling the computation and for comparison to other tests.

The F-tests indicate that both site-to-site components and sample-within-site components are significant. This is in contrast to the

TABLE 3

SUMMARY OF STATISTICAL ANALYSES

(SITE TO SITE ANALYSIS)

	Air Content, Pressure			Air Content, Chace			Slump			Unit Weight		
	Site 1	Site 2	Site 3	Site 1	Site 2	Site 3	Site 1	Site 2	Site 3	Site 1	Site 2	Site 3
Between Samples	1.1345	2.5686	0.8116	2.6093	2.0215	0.6044	4.0191	1.8208	3.1446	3.4949	4.6644	2.4261
Within Samples	0.1387	0.0558	0.0419	0.4775	0.2356	0.1962	0.2262	0.1133	0.0704	1.6434	1.3496	0.4437

previously discussed pressure meter results where the site-to-site components were not significant. The standard deviations computed for the Chace test are 0.11 percent for the site means and 0.32 percent for the sample means. It may be noted that the standard deviation for the Chace meter sample means is twice that of the pressure meter. Again it is pointed out that air contents by Chace meter are determined to the nearest one-half percent in the field and that the Chace test might well be used as an indicator of the relative air content but not as a test to determine the precise air content. The sample-within-site standard deviation is 0.69 percent which is the same as the pressure meter.

A histogram showing the distribution of air content by the Chace meter for all sites is presented in Figure 2. The values plotted are sample means. This distribution does approach a normal distribution, but an interesting observation may be made. The figure shows three distinct small peaks. These peaks occur at the mean Chace air content for each site or if one were to locate the means of each site on Figure 2, they would fall at each peak. This does not happen in the case of pressure meter results as Figure 1 clearly shows. The pressure meter distribution is nearer to a normal distribution. The distribution for Chace is more disperse, thus showing its higher variability as indicated by the higher standard deviation calculated for sample means.

From Table 2 the site to site standard deviation is 0.3%. Confidence limits placed on the site mean indicate that there is a confidence of 95% that the site mean lies between $\bar{X}_{site}$ + 0.2% and $\bar{X}_{site}$ - 0.2%. Also, the 95% confidence limits on a sample mean is $\bar{X}_{sample} \pm 0.6\%$. This last figure is interesting when it is compared to the pressure meter results. In the analysis of the pressure meter data 95% confidence limits were determined to be $\bar{X}_{sample} \pm 0.3\%$ (Table 1). This, once again indicates

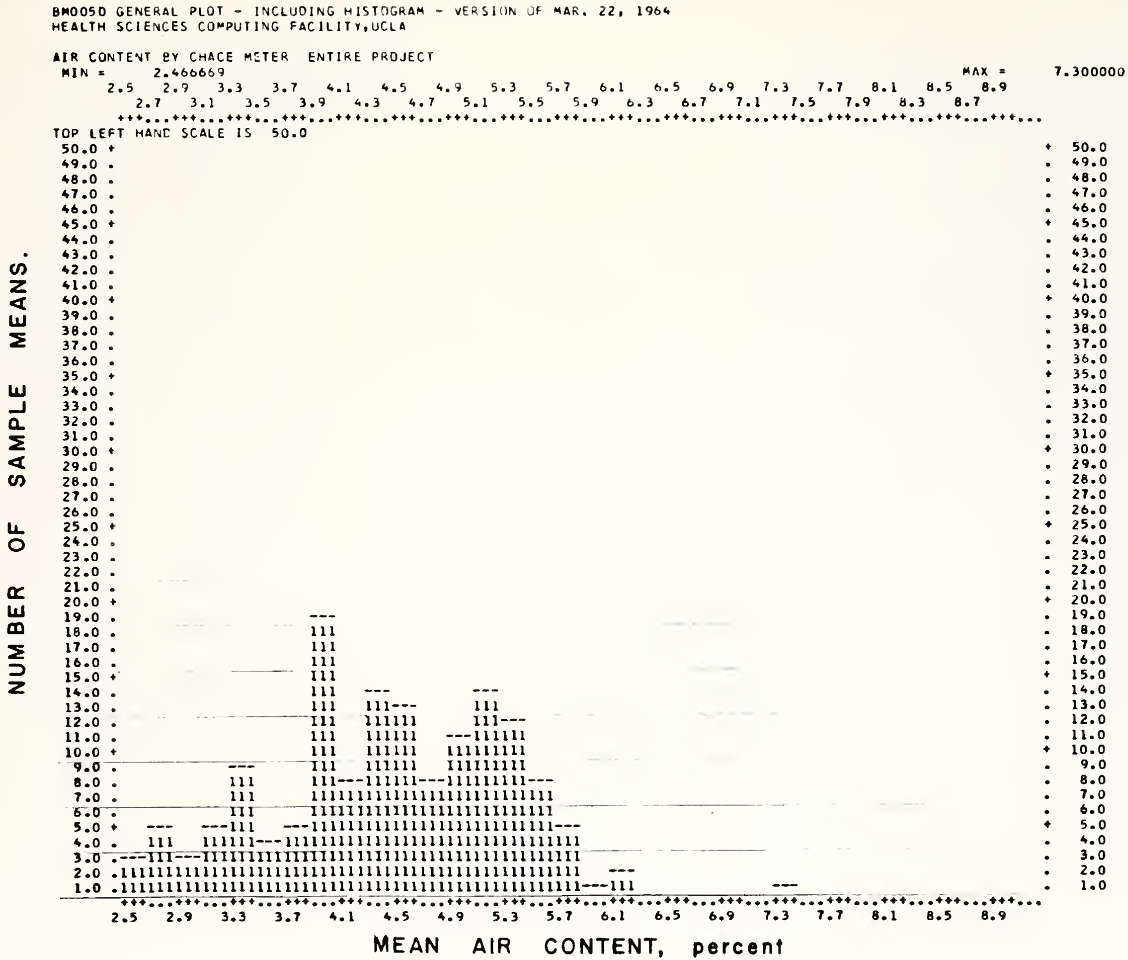

FIG. 2 HISTOGRAM FOR THE DISTRIBUTION OF AIR CONTENT MEASURED BY CHACE METER, ALL SITES.

the pressure air content test to be statistically more reliable than the Chace test.

If one were to compare the three sites in an effort to check dispersion of data, Site 3 stands out as being more consistent than the other two sites. This is true because there were few adjustments made in the air entraining agent and also less changing of the water content. Site 2 shows a sort of "sinusoidal" shape indicating trends which were not immediate but occurred over a number of samples. A plot of the pressure air content data also substantiates this. Site 2 was a central mix project and this operation had difficulties with its air dispenser which resulted in the distribution indicated. Site 2 also has the greatest amount of dispersion of the three sites.

As in the pressure meter analysis, a one-way ANOV was conducted, and the results are summarized in Table 3. Again observable differences occur in the MS and standard deviation terms from site to site. As in the pressure method analysis, the within Sample Means Square term for Site 1 is at least twice that of Sites 2 and 3 which are very nearly equal.

Slump Test

The ANOV for the slump test is similar to that in Table 1. The sources of variation (site-to-site variation, sample-within-site and error terms) are the same used for the two air content tests. Table 2 gives a summary of the statistical analysis of the slump phase of this investigation for the factorial model.

The F-test indicates that at a 0.05 α-level the site-to-site variation is not significant but the sample-within-site variation is. This is what would be expected in light of the characteristics of the slump test. The slump test is a measure of water content and therefore will vary as the water content varies. The more one changes the adjustment on the water indicator of a mixer the more the slump should change. In the light of this, one would expect Site 2, the central mix project, to show the least variation in slump which it does. Both the dual-drum paver sites show more spread in slump than Site 2. In the central mix operation there were relatively few changes in water content compared to the operations using dual-drum pavers.

The distribution of slump for all sites is presented in Figure 3. The values therein plotted are sample means. The histogram shows a close grouping of data which is a tight, almost normal, distribution. The overall mean of the slump is, for all practical purposes, three inches. There is a slight tendency for each site to approximate a normal distribution which becomes more pronounced when all three sites are lumped in Figure 3. The histogram for Site 2 is tighter than those for Sites 1 and 3 which substantiates what was said above concerning the central mix plant.

The 95% confidence limits on the site mean are $\pm$ 0.3% while 95% confidence limits on the sample mean are $\pm$ 0.4%. Site 2 had the smallest range in slump values, i.e., it exhibited both the highest minimum and lowest maximum slump.

As in the previous analyses, a one-way ANOV was performed on the slump data for each site and a summary of these results are presented in Table 3. Note that the between sample standard deviation is lowest for Site 2 bearing out the observation made from the factorial analysis that the variances for Site 2 were smaller, i.e. Site 2 exhibited better control as far as slump measurements were concerned.

Unit Weight

The distribution of unit weight from all sites is presented in Figure 4. As with the other three test methods, sites, sample-within-site

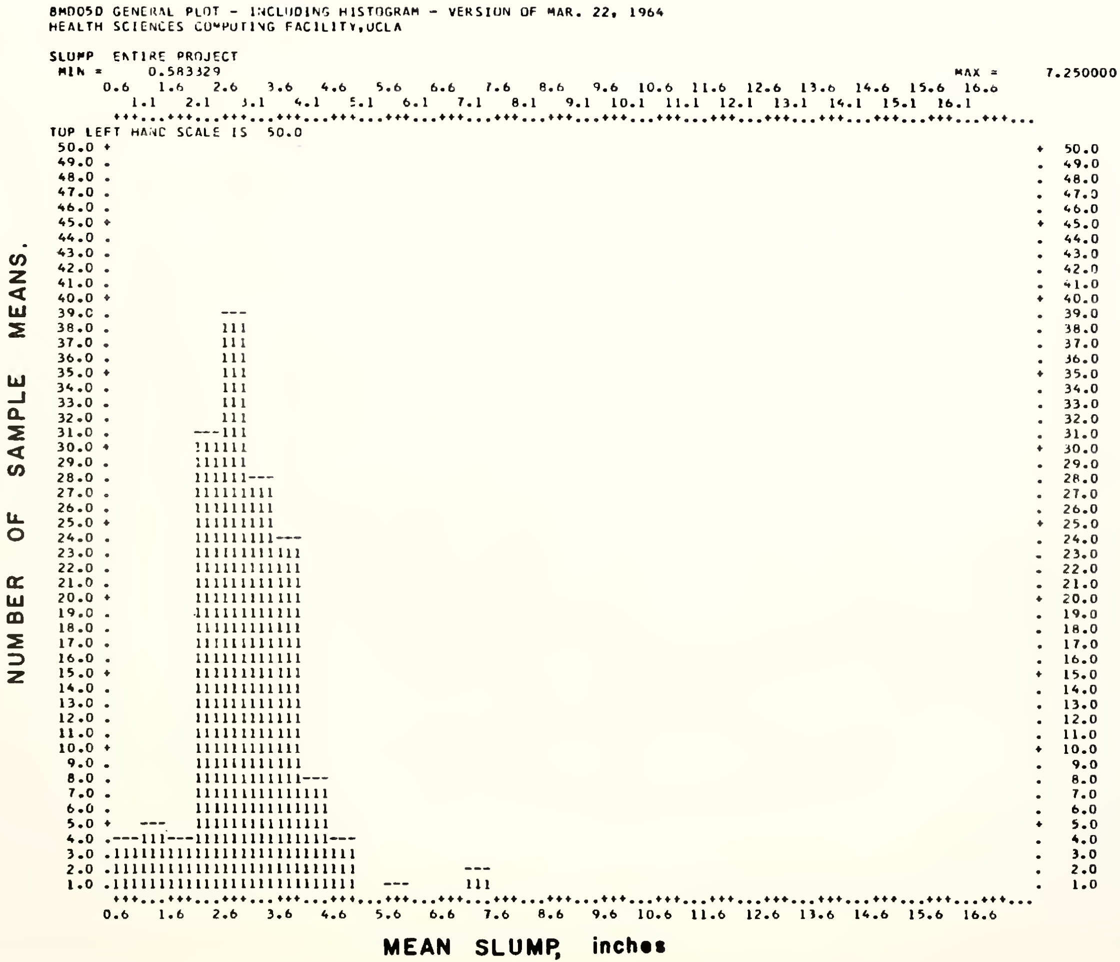

FIG. 3 HISTOGRAM FOR THE DISTRIBUTION OF SLUMP, ALL SITES.

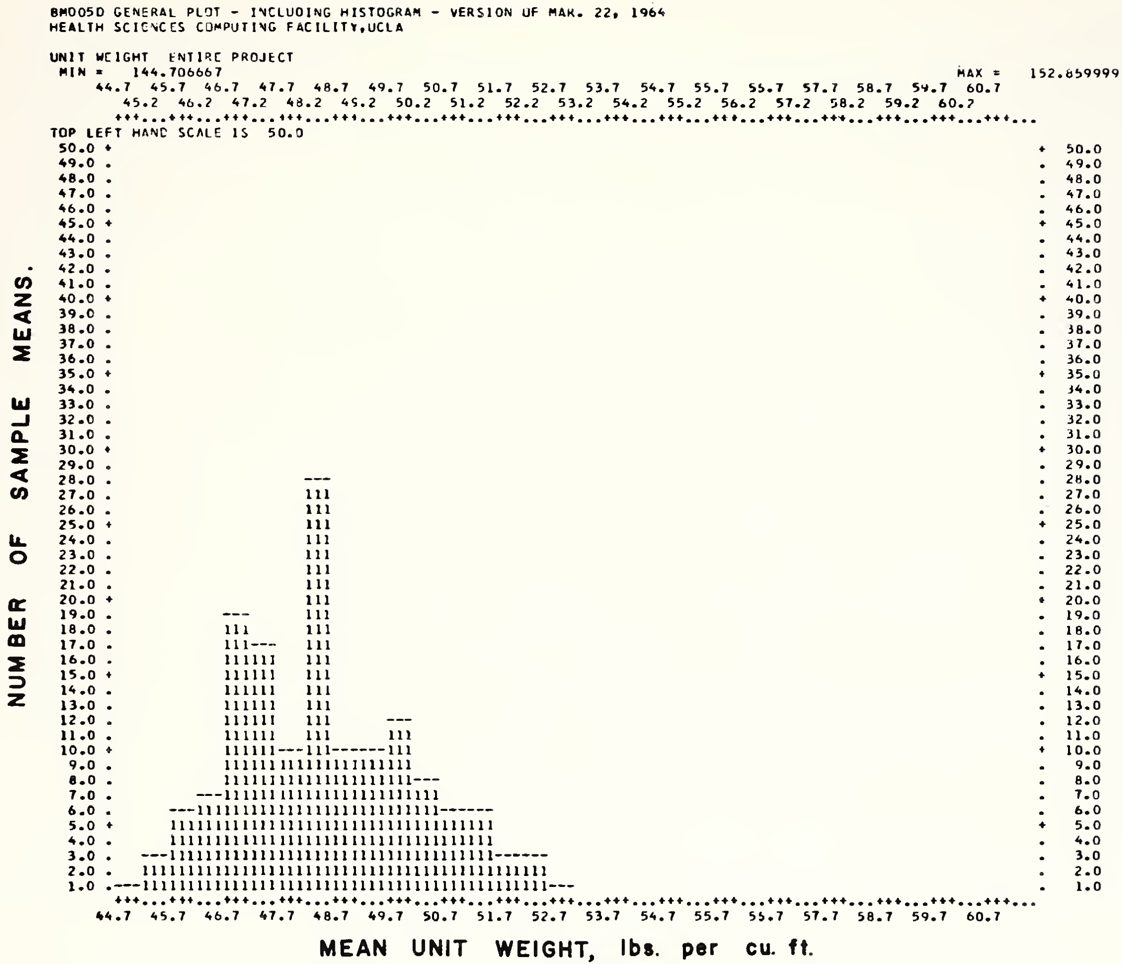

FIG. 4 HISTOGRAM FOR THE DISTRIBUTION OF UNIT WEIGHT, ALL SITES.

and the error term were the components of variation. Noting the site means and comparing these with the histogram it can be seen that the three peaks in the overall distribution correspond very closely to the three site means. Evidently changes in materials from site to site cause a definite and obvious shift in the individual site distributions that is reflected in the overall distribution.

A summary of the results from the statistical analysis is presented in Table 2. From the ANOV it was determined by F-tests that both the site component and the sample-within-site component are significant. The site component is highly significant as would be expected since from site-to-site the aggregate used varied in specific gravity and the unit weight reflected this change.

The observed error term (Table 2) is 1.15 lbs. indicating that a unit weight determination can have an error of 1.15 lbs. due to chance alone. The 95% confidence limits on the sample mean are $\pm$ 1.2 lbs. (i.e. 95% confident that the true mean lies between $\bar{X}_{sample} \pm 1.2$ lbs.). This shows that there is a great deal of variability involved in the performance of this test. This wide range might be due to variation of air content, water content of concrete or the amount of stiffness allowed to occur before testing. The longer the concrete is allowed to set, the more difficult it will be to compact it into the yield bucket. This also may lead to large voids of entrapped air in the stiffening concrete.

As in the analysis of the other three test methods, a one-way ANOV was performed on the unit weight data and a summary of the results are tabulated in Table 3. Site 2 exhibits a greater variability than do Sites 1 and 3. This is consistent with the observations made on the results of the analysis of air content data and is what would be expected since variations in air content cause the unit weight to vary accordingly.

Correlations

With the amount of data available and since the tests for air, slump and unit weight were made on the same sample it was considered advantageous to obtain information regarding correlations between the tests. Table 4 (see Appendix) presents a summary of this work. Significant correlations were found between the pressure meter air content test and the Chace meter air content test as well as with unit weight. Since both the pressure meter and the Chace meter measure air content and the air content influences the unit weight of this concrete, these significant correlations were expected. Also, there was a correlation between air content measured by the pressure test and slump, however, the correlation coefficient is not large. The correlation between air content by Chace meter and slump is not significant.

The correlation coefficients presented are the "r" values and even though significant correlations do exist there is a large amount of scatter. The predictability is relatively poor in a number of the correlations.

The correlation between air content measured by the Chace meter and unit weight is highly significant. This is in agreement with the significant correlation between air content by the pressure meter and unit weight previously noted. The correlation coefficients are negative indicating that as air content increases unit weight decreases. Both Chace air content vs. slump and slump vs. unit weight are not significant. See Table 5 in the Appendix for tabulation of confidence limits on the correlation coefficients.

Differences Between Replicate Observations

As mentioned before, a time dependency was observed when the air content was measured by the pressure meter. As a result, an analysis of the difference between replicate observations was performed. Table 6 presents a summary of this analysis. The differences between replicate 1 and replicate 2 is significant at the 0.05 α-level for all three sites. This is also true for the difference between replicate 1 and replicate 3. Replicate 2 and replicate 3 difference are not significant except in the case of Site 1 where the results are extremely close to the borderline. These results indicate that signal change in air content occurred between the first and second replicate.

As a consequence of this finding, correlation analyses of the third pressure reading versus the mean of the Chace meter was made. The mean of the Chace was used since these air contents were taken immediately after the third pressure reading and the time involved for three Chace readings is small. The correlation coefficients for each site and over all three sites are shown in Table 7. A comparison of these coefficients with those of the mean pressure versus the mean Chace show that a general trend to a lower coefficient for the case of third pressure versus the mean of the Chace meter reading. Considering the results of the analysis of differences, a higher correlation could be expected. One possible answer to the apparent contradiction is that the Chace meter air contents are measured to only the nearest one-half percent while the pressure meter readings are to the nearest one-tenth percent. A more realistic comparison might be to round the pressure meter readings to the nearest one-half percent and then make the analysis.

Basically the analysis of the differences indicates statistically significant changes in air content measured with the pressure meter as a function of time. However, the correlation of the third pressure meter reading with the mean of the Chace meter readings is inconclusive in this aspect of the analysis.

QUALITY CONTROL APPLICATIONS

It is important to understand that a quality control system depends upon the data used to establish the system. Control procedures therefore are no better than the data used to establish them and it is obviously necessary to obtain this data in some manner. There are two approaches to this problem. One approach is to rely on past data, data collected by examining records of construction, etc. The other approach sets out to obtain the data required via a preliminary testing program.

There are several problems associated with using past data. One of the most obvious is lack of reliability. The possibility is always present that only test results that met specifications were recorded. This situation may not arise out of desire to falsify records but rather from a conscientious effort to maintain good control in the field. For example, a situation may arise when something in the manufacturing process goes awry, an acceptance test is made which detects the error and appropriate steps are taken to correct the situation following which another test on the product is made and recorded. The testing has served its purpose, an error was detected and corrected, but only the last test result recorded.

For statistical evaluation of the process, the out-of-specification result is just as important as the within specification result if a realistic estimate is to be made of the variation. For this reason the second method of obtaining the so called historical, or past data, is used when there is a scarcity of information or there is reason to suspect the past data. This investigation is of the second type and operated independent of acceptance sampling.

It should be noted at this point that there are certain limitations associated with the results of this investigation. Only one operator and one piece of testing equipment were utilized for each test method conducted. There is, therefore, no estimate available of operator or equipment variability. It is a recognized fact that these variables may be significant. Another limitation arises from the fact that only three sites were checked and these were all interstate-type construction.

In the preceding section entitled Analysis of Data, the measures of central tendency and components of variability have been presented. The problem is to now apply these results to establish a realistic quality control program that may be implemented and used in the field.

The typical data plot in Figure 5 shows the fluctuation of the sample means. The variability of the product, plastic portland cement concrete, is represented by these fluctuations. One method of quality control is to establish control limits based on the data at hand and to use these limits to "control the quality" on future jobs. It is of no practical value to place the calculated limits on the data plots of the sites investigated since the calculated limits are based on the measured variability of these sites and therefore practically all of the data would fall within these limits.

For purposes of illustration, a variation that is considered to be reasonable from analysis of the data will be used and the use of control limits demonstrated in the following pages. A point should be made here concerning the distribution of the sample means. It is possible that the population of sample means is not normally distributed and normality is one of the assumptions underlying the concepts of control limits. If subgroups of 4 or 5 are used, the central limit theorem comes into play and the normalization effects is fairly strong. It is therefore better at times to use "moving means" in constructing control charts.

There are basically three types of control charts that are of use in the application of statistical quality control to the manufacture of fresh portland cement concrete. These charts are the $\overline{X}$-charts, R-chart and the σ-chart. All three of these charts provide a graphic representation of variation from point to point (i.e., sample to sample). An objective of using one or a combination of these charts is to keep track of the process so that some type of corrective action may be taken whenever the process goes "out of control" or a trend toward the control limits, indicating the possibility that an assignable cause is adding to the variation.

In concept, the control limits form a band within which fluctuations in the measured values are due to random or chance variation in the process. Observations which fall outside these limits more than a predetermined percentage of the time cannot be explained by chance causes alone and hence must be due to an assignable cause or a change occurring in the process. For example, having estimates of the components of variability associated with air content determinations, control limits may be computed and a control chart drawn. The air contents are plotted on the chart as the samples are tested during the manufacturing of the portland cement concrete. As the process proceeds, it may be noted that the air contents begin to decrease and fall outside the lower limit, hence, some assignable cause should be responsible for this change. A check of the process may show a defective dispenser, a change in sand gradation or some other recognizable cause that has resulted in the process going out of control. When this cause has been identified and corrective action taken, the process should again come into control.

If specifications have been written so that maximum and minimum values are given which form a band narrower than control limits based on the inherent variability of the process, it will be impossible to manufacture a product that will be within the specification all of the time (the percentage outside will naturally depend upon specification limits and the known standard deviation).

To illustrate one use of control limits, moving means have been computed for the data and a plot is shown in Figure 5. The moving means are averages of three sample means. The means of Samples 1, 2 and 3 are averaged and this is the first "moving mean." Then sample means 2, 3 and 4 are averaged and this is the second moving mean. This is continued for sample means 3, 4 and 5, etc., and a plot of the "moving mean" is obtained.

Assumed values used in the determination of control limits are based on the one-way ANOV and considerations of what is reasonable to expect based on field experience. The limits are for 3-σ control limits which would include approximately 99.7 percent of data if a job were operating in control. Note that even if a job were operating in control, 0.3% of the data could fall outside the control limits due to chance variation alone. If the limits were based on a 0.05 α-level, then 95% of the data would fall within the limits in the long run and 5% could fall outside the limits due to chance variation alone. This illustrates the point that because one or two observations fall outside the control limits does not necessarily mean process has gone "out of control."

Assuming a $\sigma_{\bar{X}}'$ of 0.60 for air content by pressure method, 3-σ control limits are: $\bar{\bar{X}} \pm 1.732\ (0.52)$ or $\bar{\bar{X}} \pm 0.90$. Applying these limits to the data plot of Figure 5 it may be noted that for Site 1 about 15% of the sample means are outside control limits hence one would conclude that some adjustments should be made. A similar plot of the data for Site 2 would show approximately 40% of data outside the limits. The job is in poor control, action should be taken. By contrast Site 3 exhibits

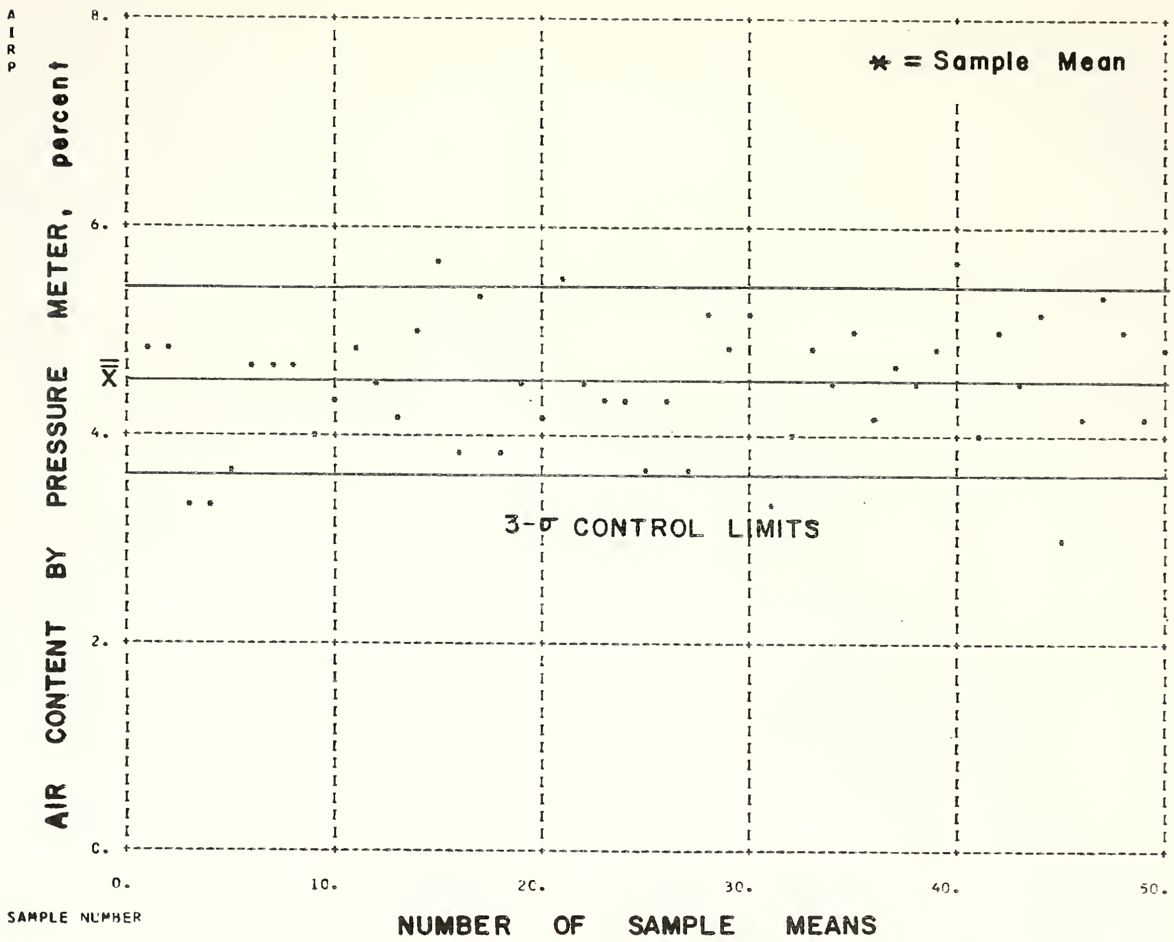

FIG. 5 DATA PLOT FOR AIR CONTENT MEASURED BY PRESSURE METER, SITE I

the best control, only 4% of the data would fall outside the limits. If the moving average concept is used, the same general conclusion may be reached and additional information concerning trends in the data may also be noted which may be valuable in field control.

The control limits determined from the assumed values of the components of variation are shown on the overlay sheets for the data plots. The 3-σ limits are in blue while the 95% limits are not shown. These limits are to be considered illustrative only since the variables of operator and equipment have not been evaluated.

With estimates of the components of variance available it is possible to take a critical look at present specifications. As mentioned previously, even though a process is "in control" if the variability of the process is high it may be incapable of producing a product always inside the specification limits. If this is the case, there are several possible avenues of action. The specifications should be examined to determine if the limits actually need to be as tight as they are. Also, the process itself should be examined to determine if any adjustments or changes are possible which will reduce the inherent variability of the process itself. This situation also points the way towards acceptance testing. A process may be operating "in control" and still have the product falling outside specifications. Operating "in control" does not insure that a product will meet specifications.

There are other ways of providing control procedures and one such method is to use tolerance limits. For example, if air content is desired to be between 4-7% and the variance is known, then a range of means may be used. If the variation on a site is known and 3-σ limits determined to be 5.5% $\pm$ 0.90%, then the average air content can be 5.5% $\pm$ 0.60% for a process in control and the material will meet the specified 4-7% air content providing the process remains in control. Another approach is to specify a mean and allow a standard deviation range. For example, specify a mean of 5.5%, the standard deviation may then be less than or equal to 0.5% for 3-σ limits and the product will pass the 4-7% specification limits. Tables can be set up for various means and various standard deviations, allowing a contractor operating with a known standard deviation a certain latitude in mean air content. The same may be accomplished by testing standard deviation and then stating that if a standard deviation of so much is occurring then the mean air content must be within certain limits for the product to meet specification limits.

APPENDIX

TABLE 4

TABULATION OF CORRELATION COEFFICIENTS

AND SIGNIFICANCE TESTS

All Sites

Variables	r	t_{148}	$t_{\alpha=0.001}$	Significance
Pressure v. Chace	0.6060	9.2675	3.29	Highly Significant
Pressure v. Slump	0.3368	4.3516	3.29	Significant
Pressure v. Unit Wt.	-0.5491	8.6351	3.29	Highly Significant
Chace v. Slump	0.1296	1.5900	3.29	Not Significant
Chace v. Unit Wt.	-0.6445	10.2540	3.29	Highly Significant
Slump v. Unit Wt.	-0.1856	2.2977	3.29	Not Significant

Pressure v. Chace by Sites

Site	r	t_{148}	$t_{\alpha=0.001}$	Significance
1	0.5130	4.1405	3.51	Significant
2	0.7288	7.3744	3.51	Highly Significant
3	0.7247	7.2861	3.51	Highly Significant

Interpretation of Correlation Coefficients[1]

Correlation Coefficient	Relationship Demonstrated
1.0	Perfect
0.9	Very good
0.8	Good
0.7	Fair
0.6	Poor
0.5 or less	Very poor

[1] Hughes, C. S., Enrick, N. L. and Dillard, J. H., "Applications of Some Statistical Techniques to Experiment in Highway Engineering", Virginia Council of Highway Investigations and Research in Cooperation with the U.S. Bureau of Public Roads, February 1964.

TABLE 5

CONFIDENCE LIMITS ON CORRELATION COEFFICIENTS, ALL SITES

	r	Limits
Pressure vs. Chace	0.6060	0.49 to 0.70
Pressure vs. Slump	0.3368	0.22 to 0.51
Pressure vs. Unit Weight	-0.5491	-0.65 to -0.42
Chace vs. Slump	0.1296	-0.04 to 0.29
Chace vs. Unit Weight	-0.6445	-0.74 to -0.52
Slump vs. Unit Weight	-0.1856	-0.33 to -0.015

TABLE 6

SUMMARY OF ANALYSIS OF DIFFERENCES

BETWEEN REPLICATE OBSERVATIONS

SITE	OBSERVATION DIFFERENCE=d	$\bar{d}$	S	$t = \frac{\bar{d}}{S}$	t_α	Significant
	$X_1 - X_2$	0.256	0.03727	6.87	2.01	Yes
1	$X_1 - X_3$	0.398	0.07365	5.40	2.01	Yes
	$X_2 - X_3$	0.142	0.07022	2.02	2.01	Yes
	$X_1 - X_2$	0.136	0.04452	3.05	2.01	Yes
2	$X_1 - X_3$	0.190	0.05049	3.76	2.01	Yes
	$X_2 - X_3$	0.054	0.03360	1.61	2.01	No
	$X_1 - X_2$	0.208	0.03486	5.97	2.01	Yes
3	$X_1 - X_3$	0.302	0.03619	6.41	2.01	Yes
	$X_2 - X_3$	0.024	0.02467	0.97	2.01	No

TABLE 7

TABULATION OF CORRELATION COEFFICIENTS, r

THIRD PRESSURE VERSUS $\overline{X}$ CHACE

Site	r
1	0.3776
2	0.7125
3	0.7015
All Sites	0.5677

BIBLIOGRAPHY

1. "Basic Statistical Methods for the Concrete Laboratory", Miscellaneous Paper No. 6-41, Corps of Engineers, U.S. Army, Waterways Experiment Station, October 1951 (Revised June 1953).

2. Bennett, C. A. and Franklin, N. L., Statistical Analysis in Chemistry and the Chemical Industry, John Wiley and Sons, Inc., 1954.

3. Burr, I. W., Engineering Statistics and Quality Control, McGraw-Hill Book Company, Inc., 1953.

4. Duncan, A. J., Quality Control and Industrial Statistics, Richard D. Irwin, Inc., 1959.

5. Grant, E. L., Statistical Quality Control, McGraw-Hill Book Company, Inc., 1946.

6. Hanna, S. J. and McLaughlin, J. F., "The Development of Precision Statements for Several ASTM Test Methods", Proc. ASTM, 1963 (in Press).

7. "Highway Quality Control", Proposal for a Research Project by the Michigan State Highway Department in Cooperation with the Bureau of Public Roads, February 1963.

8. "Highway Quality Control and Research", Fourth Annual Highway Conference, Michigan College of Mining and Technology, October 3 and 4, 1963.

9. Hughes, C. S., Enrick, N. L. and Dillard, J. H., "Applications of Some Statistical Techniques to Experiment in Highway Engineering", Virginia Council of Highway Investigations and Research in Cooperation with the U.S. Bureau of Public Roads, February 1964.

10. Irick, Paul, "Basic Concepts of Statistical Quality Control", Proceedings of 50th Annual Purdue Road School, March 1964.

11. "Research Project Prospectus for an Investigation of Improved Quality Control in Highway Construction", Illinois Research Suggestion No. 85, Illinois Division of Highways, Bureau of Research and Planning, March 19, 1963.

12. "Statistical Methods for Quality Control of Road and Paving Materials", ASTM Special Technical Publication No. 362, June 1963.